

A note on Burer’s copositive representation of mixed-binary QPs

Immanuel M. Bomze · Florian Jarre

Received: 27 September 2009 / Accepted: 7 January 2010 / Published online: 23 January 2010
© Springer-Verlag 2010

Abstract In an important paper, Burer (Math. Program Ser. A 120:479–495, 2009) recently showed how to reformulate general mixed-binary quadratic optimization problems (QPs) into copositive programs where a linear functional is minimized over a linearly constrained subset of the cone of completely positive matrices. In this note we interpret the implication from a topological point of view, showing that the Minkowski sum of the lifted feasible set and the lifted recession cone gives exactly the closure of the former. We also discuss why feasibility of the copositive program implies feasibility of the original mixed-binary QP, which can be derived from the arguments in Burer (Math. Program Ser. A 120:479–495, 2009) without any further condition.

Keyword Copositive programming

1 Introduction and preliminaries

In an important paper, Burer [2] recently showed how to reformulate general mixed-binary quadratic optimization problems (QPs) into copositive programs where a linear functional is minimized over a linearly constrained subset of the cone of completely positive matrices. By this method, the complexity of a possibly non-convex objective in the QP is shifted entirely to the feasibility question whether or not a symmetric matrix Y is completely positive, i.e., whether it admits a factorization $Y = FF^T$ where F is a possibly rectangular matrix with no negative entries. In order to establish

I. M. Bomze (✉)
University of Vienna, Vienna, Austria
e-mail: immanuel.bomze@univie.ac.at

F. Jarre
University of Düsseldorf, Düsseldorf, Germany

full equivalence of the programs instead of having just another conic relaxation, [2] employs the following assumption: if we relax the binarity condition on a variable x_j in the usual linear relaxation way, namely $0 \leq x_j \leq 1$, then these conditions are already implied by the (other) linear constraints on \mathbf{x} .

Let a symmetric matrix $Q \in \mathbb{R}^{n \times n}$, a vector $\mathbf{c} \in \mathbb{R}^n$, a matrix

$$A = \begin{bmatrix} \mathbf{a}_1^\top \\ \vdots \\ \mathbf{a}_m^\top \end{bmatrix} \in \mathbb{R}^{m \times n},$$

a vector $\mathbf{b} \in \mathbb{R}^m$, and a subset $B \subseteq \{1, \dots, n\}$ be given.

We consider mixed binary quadratic programs of the form

$$(P) \min \left\{ f(\mathbf{x}) := \mathbf{x}^\top Q \mathbf{x} + 2\mathbf{c}^\top \mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \in \mathbb{R}_+^n, \mathbf{x}_B \in \{0, 1\}^{|B|} \right\}.$$

Note that the polyhedron $\{\mathbf{x} \in \mathbb{R}_+^n : A\mathbf{x} = \mathbf{b}\}$ almost coincides with the linearly relaxed feasible set of (P); only the constraints $x_j \leq 1$ for $j \in B$ are missing. Burer [2] defines a *key condition*, requiring the following implication:

$$A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0} \implies x_j \leq 1 \text{ for all } j \in B. \tag{1}$$

In other words, the constraints $x_j \leq 1$ are redundant for the linear relaxation. In [2] it is also pointed out that the key condition can easily be satisfied by introducing $|B|$ additional slack variables $s_j, j \in B$.

Let \mathcal{C}^* denote the cone of completely positive symmetric $(n + 1) \times (n + 1)$ matrices, see e.g. [1]. The completely positive relaxation of (P) is then given by (C)

$$\min \left\{ Q \bullet X + 2\mathbf{c}^\top \mathbf{x} : (\mathbf{x}, X) \in \mathbb{R}^n \times \mathbb{R}^{n \times n}, \begin{bmatrix} 1 & \mathbf{x}^\top \\ \mathbf{x} & X \end{bmatrix} \in \mathcal{C}^*, \right. \\ \left. \mathbf{a}_i^\top \mathbf{x} = b_i, \mathbf{a}_i^\top X \mathbf{a}_i = b_i^2 \ (1 \leq i \leq m), x_j = X_{jj} \ (j \in B) \right\}.$$

We denote the feasible sets of (P) and (C) by $\text{feas}(P)$ and $\text{feas}(C)$, and

$$\text{feas}^+(P) := \text{conv} \left(\left\{ \begin{bmatrix} 1 & \mathbf{x}^\top \\ \mathbf{x} & \mathbf{x}\mathbf{x}^\top \end{bmatrix} : \mathbf{x} \in \text{feas}(P) \right\} \right), \\ \text{feas}^+(C) := \left\{ \begin{bmatrix} 1 & \mathbf{x}^\top \\ \mathbf{x} & X \end{bmatrix} : (\mathbf{x}, X) \in \text{feas}(C) \right\}.$$

While $\text{feas}^+(C)$ is just a linear reparametrization of $\text{feas}(C)$, the relation between the sets $\text{feas}^+(P)$ and $\text{feas}(P)$ is a bit more involved. Anyhow, it is evident by convexity that

$$\text{feas}^+(P) \subseteq \text{feas}^+(C). \tag{2}$$

In [2] it is beforehand assumed that (P) is feasible. Obviously this implies feasibility of (C). Somehow less explicit, the proofs in [2] also establish the converse, so that

full equivalence of the two programs (P) and (C) is established. Since it may not be immediate that the arguments in [2] do not need any of the conditions (1) or (4) below, we provide an alternative, probably more explicit argument:

Suppose $(\hat{\mathbf{x}}, \hat{X}) \in \text{feas}(\text{C})$. Then $\hat{\mathbf{x}} \in M = \{\mathbf{x} \in \mathbb{R}_+^n : A\mathbf{x} = \mathbf{b}\}$, so $M \neq \emptyset$. If $B = \emptyset$, we are done. Else, for $j \in B$ and $(\mathbf{x}, X) \in \text{feas}(\text{C})$ the determinant of $\begin{bmatrix} 1 & x_j \\ x_j & X_{jj} \end{bmatrix}$ is not negative since $\begin{bmatrix} 1 & \mathbf{x}^\top \\ \mathbf{x} & X \end{bmatrix}$ is positive-semidefinite. Thus $x_j = X_{jj} \geq x_j^2$, and $x_j \geq 0$ implies $x_j \leq 1$ for $j \in B$. Therefore we may and do add the constraints $x_j \leq 1$ for $j \in B$ to both (P) and (C), without changing the feasibility question as obviously also

$$M' := \{\mathbf{x} \in M : x_j \leq 1 \text{ for all } j \in B\} \supseteq \text{feas}(\text{P}).$$

Next denote by \mathbf{e}_j the j th column of the $n \times n$ identity matrix and define $S_B := -\sum_{j \in B} \mathbf{e}_j \mathbf{e}_j^\top$ as well as $\mathbf{e}_B := \frac{1}{2} \sum_{j \in B} \mathbf{e}_j \in \mathbb{R}^n$. Then

$$f_B(\mathbf{x}) := \mathbf{x}^\top S_B \mathbf{x} + 2\mathbf{e}_B^\top \mathbf{x} = \sum_{j \in B} x_j(1 - x_j) \geq 0 \text{ for all } \mathbf{x} \in M'$$

with equality if and only if $x_j \in \{0, 1\}$ for all $j \in B$. So the QP (without binary variables) $\gamma_B := \min \{f_B(\mathbf{x}) : \mathbf{x} \in M'\}$ satisfies (1) since this condition is void, and Burer’s result [2, Theorem 2.6] shows that γ_B equals the minimal value of the respective copositive program

$$\gamma_B = \min \left\{ 2\mathbf{e}_B^\top \mathbf{x} + S_B \bullet X : (\mathbf{x}, X) \in \text{feas}(\text{C}), x_j \leq 1 (j \in B) \right\}.$$

However, as established above, $(\hat{\mathbf{x}}, \hat{X})$ is feasible also to this copositive program, hence $\gamma_B \leq \sum_{j \in B} \hat{x}_j - \sum_{j \in B} \hat{X}_{jj} = 0$. On the other hand, we know $\gamma_B = \min \{f_B(\mathbf{x}) : \mathbf{x} \in M'\} \geq 0$ by construction. So $\gamma_B = 0 = f_B(\bar{\mathbf{x}})$ for some $\bar{\mathbf{x}} \in M'$, and therefore $\bar{\mathbf{x}} \in \text{feas}(\text{P})$.

In the sequel, the study of the recession cone

$$L_\infty := \{\mathbf{d} \in \mathbb{R}_+^n : A\mathbf{d} = \mathbf{o}\}, \quad \text{and} \quad L_\infty^+ := \text{conv} \left(\left\{ \begin{bmatrix} 0 & \mathbf{o}^\top \\ \mathbf{o} & \mathbf{d}\mathbf{d}^\top \end{bmatrix} : \mathbf{d} \in L_\infty \right\} \right)$$

will be important. Burer shows in [2] that the key condition (1) implies

$$\text{feas}^+(\text{C}) = \text{feas}^+(\text{P}) + L_\infty^+. \tag{3}$$

To complement this result we define a *weak key condition*:

$$A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{o} \quad \implies \quad x_j \text{ are bounded for all } j \in B. \tag{4}$$

and show that under the weak key condition the set on the right-hand side in (3) equals the closure $\overline{\text{feas}^+(\text{P})}$ of $\text{feas}^+(\text{P})$:

$$\overline{\text{feas}^+(\text{P})} = \text{feas}^+(\text{P}) + L_\infty^+. \tag{5}$$

By the same arguments that led to [2, Property (2)], we observe that

$$\text{under assumption (4), we have } d_j = 0 \text{ for all } j \in B \text{ and } \mathbf{d} \in L_\infty. \tag{6}$$

Indeed, choose any $\mathbf{x} \in \text{feas}(\text{P}) \neq \emptyset$ and any $\mathbf{d} \in L_\infty$. Then $A(\mathbf{x} + \gamma\mathbf{d}) = \mathbf{b}$ and $\mathbf{x} + \gamma\mathbf{d} \in \mathbb{R}_+^n$ for all $\gamma > 0$, so the weak key condition (4) implies for $j \in B$ that $x_j + \gamma d_j \leq K$ for some $K > 0$. Since this holds for arbitrarily large $\gamma > 0$, we obtain $d_j = 0$.

The differences between (1) and (4) may become more apparent when realizing that the latter can be enforced by adding a single slack variable $s \geq 0$ and a single linear constraint of the form

$$\sum_{j \in B} x_j + s = |B|,$$

rather than introducing $|B|$ slack variables to enforce (1). Already for medium-sized conic optimization problems, the number of variables and constraints become a crucial problem characteristic from the view of performance of interior-point algorithms, so that results like (4) \implies (5) may be of interest.

2 The results

Proposition 2.1 *If (P) is feasible, the weak key condition (4) implies*

$$\overline{\text{feas}^+(\text{P})} = \text{feas}^+(\text{P}) + L_\infty^+.$$

Proof Recall from [3] that a convex polyhedron not containing a straight line can be written as the convex combination of its extreme points and its extreme rays. Now the extreme rays of $\text{feas}(\text{P})$ are contained in L_∞ independently of the choice of $x_j \in \{0, 1\}$ for $j \in B$. For a fixed choice of $\mathbf{v} \in \{0, 1\}^{|B|}$ the set $\text{feas}(\text{P}) \cap \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x}_B = \mathbf{v}\}$ is a polyhedron, the extreme points of which are bounded. Thus, the union of the extreme points of the finitely many polyhedra obtained for all possible selections of $\mathbf{x}_B \in \{0, 1\}^{|B|}$ is also bounded: there is a constant $M > 0$ such that $\|\mathbf{z}\| \leq M$ for all such extreme points \mathbf{z} . Hence, any $\mathbf{x} \in \text{feas}(\text{P})$ can be written as $\mathbf{z} + \alpha\mathbf{d}$ where $\|\mathbf{z}\| \leq M$ and $\mathbf{d} \in L_\infty$ with $\|\mathbf{d}\| = 1$ and $\alpha \geq 0$. Next assume that

$$\bar{Y} := \begin{bmatrix} 1 & \bar{\mathbf{x}}^\top \\ \bar{\mathbf{x}} & \bar{X} \end{bmatrix} \in \overline{\text{feas}^+(\text{P})}.$$

Then, there exists a sequence

$$Y_k := \begin{bmatrix} 1 & \mathbf{x}_k^\top \\ \mathbf{x}_k & X_k \end{bmatrix} \subseteq \text{feas}^+(\text{P}), \quad k \in \mathbb{N},$$

with $Y_k \rightarrow \bar{Y}$. By Caratheodory’s theorem, each Y_k can be written as a convex combination of at most n^2 (in fact even fewer) rank-one-matrices

$$Y_k = \sum_{r=1}^{n^2} \lambda_{k,r} \begin{bmatrix} 1 \\ \mathbf{x}_{k,r} \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{x}_{k,r} \end{bmatrix}^\top$$

with $\lambda_{k,r} \geq 0$, $\sum_r \lambda_{k,r} = 1$, and $\mathbf{x}_{k,r} \in \text{feas}(\text{P})$. By the preceding considerations, each $\mathbf{x}_{k,r}$ can be written as

$$\mathbf{x}_{k,r} = \mathbf{z}_{k,r} + \alpha_{k,r} \mathbf{d}_{k,r}$$

with $\|\mathbf{z}_{k,r}\| \leq M$, $\mathbf{d}_{k,r} \in L_\infty$, $\|\mathbf{d}_{k,r}\| = 1$, and $\alpha_{k,r} \geq 0$. The sequence

$$\left[\mathbf{z}_{k,1}; \mathbf{z}_{k,2}; \dots; \mathbf{z}_{k,n^2}; \mathbf{d}_{k,1}; \mathbf{d}_{k,2}; \dots; \mathbf{d}_{k,n^2} \right]_{k \in \mathbb{N}}$$

in \mathbb{R}^{2n^3} is bounded, and thus, it has accumulation points. By extracting a suitable subsequence, we may and do assume without loss of generality that

$$\mathbf{z}_{k,r} \rightarrow \bar{\mathbf{z}}_r, \quad \mathbf{d}_{k,r} \rightarrow \bar{\mathbf{d}}_r \quad \text{for } 1 \leq r \leq n^2.$$

Note that

$$\begin{aligned} Y_k &= \sum_{r=1}^{n^2} \lambda_{k,r} \begin{bmatrix} 1 \\ \mathbf{z}_{k,r} + \alpha_{k,r} \mathbf{d}_{k,r} \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{z}_{k,r} + \alpha_{k,r} \mathbf{d}_{k,r} \end{bmatrix}^\top \\ &= \sum_{r=1}^{n^2} \lambda_{k,r} \left\{ \begin{bmatrix} 1 \\ \mathbf{z}_{k,r} \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{z}_{k,r} \end{bmatrix}^\top + \lambda_{k,r} \alpha_{k,r}^2 \begin{bmatrix} 0 \\ \mathbf{d}_{k,r} \end{bmatrix} \begin{bmatrix} 0 \\ \mathbf{d}_{k,r} \end{bmatrix}^\top \right. \\ &\quad \left. + \lambda_{k,r} \alpha_{k,r} \left(\begin{bmatrix} 1 \\ \mathbf{z}_{k,r} \end{bmatrix} \begin{bmatrix} 0 \\ \mathbf{d}_{k,r} \end{bmatrix}^\top + \begin{bmatrix} 0 \\ \mathbf{d}_{k,r} \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{z}_{k,r} \end{bmatrix}^\top \right) \right\}. \end{aligned}$$

From $\|\mathbf{d}_{k,r}\| = 1$, nonnegativity of $\mathbf{d}_{k,r}$, $\mathbf{z}_{k,r}$, and boundedness of Y_k we deduce that the products $\lambda_{k,r} \alpha_{k,r}^2$ are bounded for $1 \leq r \leq n^2$. By considering a subsequence again, if necessary, we may therefore assume that also

$$\lambda_{k,r} \rightarrow \bar{\lambda}_r \quad \text{and} \quad \lambda_{k,r} \alpha_{k,r}^2 \rightarrow \beta_r \quad \text{for } 1 \leq r \leq n^2.$$

This also implies

$$\lambda_{k,r} \alpha_{k,r} = \sqrt{\lambda_{k,r} (\lambda_{k,r} \alpha_{k,r}^2)} \rightarrow \sqrt{\bar{\lambda}_r \beta_r} \quad \text{for } 1 \leq r \leq n^2.$$

Note that $\bar{\lambda}_r \geq 0$, and $\sum_r \bar{\lambda}_r = 1$. Taking the limit in the above decomposition of Y_k we obtain

$$\begin{aligned} Y_k &\rightarrow \sum_{r=1}^{n^2} \left\{ \bar{\lambda}_r \begin{bmatrix} 1 \\ \bar{\mathbf{z}}^r \end{bmatrix} \begin{bmatrix} 1 \\ \bar{\mathbf{z}}^r \end{bmatrix}^\top + \beta_r \begin{bmatrix} 0 \\ \bar{\mathbf{d}}^r \end{bmatrix} \begin{bmatrix} 0 \\ \bar{\mathbf{d}}^r \end{bmatrix}^\top \right. \\ &\quad \left. + \sqrt{\bar{\lambda}_r \beta_r} \left(\begin{bmatrix} 1 \\ \bar{\mathbf{z}}^r \end{bmatrix} \begin{bmatrix} 0 \\ \bar{\mathbf{d}}^r \end{bmatrix}^\top + \begin{bmatrix} 0 \\ \bar{\mathbf{d}}^r \end{bmatrix} \begin{bmatrix} 1 \\ \bar{\mathbf{z}}^r \end{bmatrix}^\top \right) \right\} \\ &= \sum_{r: \bar{\lambda}_r > 0} \bar{\lambda}_r \begin{bmatrix} \bar{\mathbf{z}}^r + \frac{1}{\sqrt{\frac{\beta_r}{\bar{\lambda}_r}}} \bar{\mathbf{d}}^r \\ \bar{\mathbf{z}}^r + \sqrt{\frac{\beta_r}{\bar{\lambda}_r}} \bar{\mathbf{d}}^r \end{bmatrix} \begin{bmatrix} \bar{\mathbf{z}}^r + \frac{1}{\sqrt{\frac{\beta_r}{\bar{\lambda}_r}}} \bar{\mathbf{d}}^r \\ \bar{\mathbf{z}}^r + \sqrt{\frac{\beta_r}{\bar{\lambda}_r}} \bar{\mathbf{d}}^r \end{bmatrix}^\top \\ &\quad + \sum_{r: \bar{\lambda}_r = 0} \beta_r \begin{bmatrix} 0 \\ \bar{\mathbf{d}}^r \end{bmatrix} \begin{bmatrix} 0 \\ \bar{\mathbf{d}}^r \end{bmatrix}^\top \\ &\in \text{feas}^+(\mathbf{P}) + L_\infty^+. \end{aligned}$$

Thus, $\overline{\text{feas}^+(\mathbf{P})} \subseteq \text{feas}^+(\mathbf{P}) + L_\infty^+$. Now consider an element $\bar{Y} \in \text{feas}^+(\mathbf{P}) + L_\infty^+$. By definition, using Caratheodory’s theorem again, and the fact that L_∞ is a cone, $\bar{Y} = \sum_{r=1}^{n^2} \lambda_r \bar{Y}_r$ is a convex combination (so $\lambda \geq \mathbf{0}$ with $\sum_r \lambda_r = 1$) of matrices \bar{Y}_r of the form

$$\bar{Y}_r = \begin{bmatrix} 1 \\ \mathbf{z}_r \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{z}_r \end{bmatrix}^\top + \begin{bmatrix} 0 \\ \mathbf{d}_r \end{bmatrix} \begin{bmatrix} 0 \\ \mathbf{d}_r \end{bmatrix}^\top$$

where $\mathbf{z}_r \in \text{feas}(\mathbf{P})$ and $\mathbf{d}_r \in L_\infty$. Hence we have proved $\bar{Y} \in \overline{\text{feas}^+(\mathbf{P})}$ once we can approximate from within $\text{feas}^+(\mathbf{P})$ every matrix of the form

$$\begin{bmatrix} 1 \\ \bar{\mathbf{z}} \end{bmatrix} \begin{bmatrix} 1 \\ \bar{\mathbf{z}} \end{bmatrix}^\top + \begin{bmatrix} 0 \\ \bar{\mathbf{d}} \end{bmatrix} \begin{bmatrix} 0 \\ \bar{\mathbf{d}} \end{bmatrix}^\top$$

where $\bar{\mathbf{z}} \in \text{feas}(\mathbf{P})$ and $\bar{\mathbf{d}} \in L_\infty$. Note that for such matrices $\bar{d}_j = 0$ for all $j \in B$ by property (6). Thus, $\bar{\mathbf{z}}_r + \alpha \bar{\mathbf{d}}_r \in \text{feas}(\mathbf{P})$ for all $\alpha \geq 0$, and therefore

$$\begin{aligned} V_r(\alpha) &:= \begin{bmatrix} 1 \\ \bar{\mathbf{z}}_r \end{bmatrix} \begin{bmatrix} 1 \\ \bar{\mathbf{z}}_r \end{bmatrix}^\top + \alpha W_r + \alpha^2 \begin{bmatrix} 0 \\ \bar{\mathbf{d}}_r \end{bmatrix} \begin{bmatrix} 0 \\ \bar{\mathbf{d}}_r \end{bmatrix}^\top \\ &= \begin{bmatrix} 1 \\ \bar{\mathbf{z}}_r + \alpha \bar{\mathbf{d}}_r \end{bmatrix} \begin{bmatrix} 1 \\ \bar{\mathbf{z}}_r + \alpha \bar{\mathbf{d}}_r \end{bmatrix}^\top \in \text{feas}^+(\mathbf{P}), \end{aligned}$$

where $W_r = \begin{bmatrix} 1 \\ \bar{\mathbf{z}}_r \end{bmatrix} \begin{bmatrix} 0 \\ \bar{\mathbf{d}}_r \end{bmatrix}^\top + \begin{bmatrix} 0 \\ \bar{\mathbf{d}}_r \end{bmatrix} \begin{bmatrix} 1 \\ \bar{\mathbf{z}}_r \end{bmatrix}^\top$. Next observe for $\alpha \geq 1$ that

$$\bar{Y}_r(\alpha) := \bar{Y}_r + \frac{1}{\alpha} W_r = \left[\left(1 - \frac{1}{\alpha^2}\right) V_r(0) + \frac{1}{\alpha^2} V_r(\alpha) \right] \in \text{feas}^+(P),$$

and by construction $\lim_{\alpha \rightarrow \infty} \sum_{r=1}^{n^2} \bar{Y}_r(\alpha) = \bar{Y}$. □

While the preceding observation sheds some light on the geometric structure of problem (P), we still need to relate the sets $\text{feas}^+(P)$ and $\text{feas}^+(C)$, to construct an optimal solution of program (P) from a solution of program (C). To this end, we establish the recession result parallel to [2, Lemma 2.7] which holds also under the weaker condition (4):

Proposition 2.2 *If (P) is feasible and the objective function is bounded over (P), the weak key condition (4) implies that Q is L_∞ -copositive, i.e.,*

$$\mathbf{d}^\top Q \mathbf{d} \geq 0 \quad \text{for all } \mathbf{d} \in L_\infty. \tag{7}$$

To be more precise, denote by $\{\mathbf{w}_1, \dots, \mathbf{w}_s\}$ the extremal points of $\text{feas}(P) \neq \emptyset$ as in the preceding proof; then boundedness of (P) is equivalent to

$$\begin{cases} \frac{1}{2} \nabla f(\mathbf{w}_r) = \mathbf{c} + Q \mathbf{w}_r \in [K_Q]^* & \text{for all } r \in \{1, \dots, s\}; \text{ and} \\ Q \text{ is } L_\infty\text{-copositive,} \end{cases}$$

where $K_Q = \{\mathbf{d} \in L_\infty : \mathbf{d}^\top Q \mathbf{d} = 0\}$ contains the recession directions of zero curvature.

Proof Choose any $\mathbf{x} = \mathbf{z} + \delta \mathbf{d} \in \text{feas}(P)$ with $\mathbf{z} \in \text{conv}(\mathbf{w}_1, \dots, \mathbf{w}_s)$, $\delta \geq 0$ and $\mathbf{d} \in L_\infty$. Then (6) implies $d_j = 0$ for all $j \in B$ so that $z_j = x_j \in \{0, 1\}$ for all $j \in B$. Hence also $\mathbf{y} = \mathbf{z} + \gamma \mathbf{d} \in \text{feas}(P)$ for all (arbitrarily large) $\gamma > 0$. Hence the relation

$$(\mathbf{z} + \gamma \mathbf{d})^\top Q (\mathbf{z} + \gamma \mathbf{d}) + 2\mathbf{c}^\top (\mathbf{z} + \gamma \mathbf{d}) = \gamma^2 \mathbf{d}^\top Q \mathbf{d} + 2\gamma \mathbf{d}^\top (\mathbf{c} + Q \mathbf{z}) + \mathbf{z}^\top Q \mathbf{z} + 2\mathbf{c}^\top \mathbf{z}$$

implies the result. □

We close by rewriting [2, Theorem 2.6] in terms of closure, to supply a topological condition by which the main result may be more transparent.

Theorem 2.1 *Under the key condition (1), we have*

$$\overline{\text{feas}^+(P)} = \text{feas}^+(C). \tag{8}$$

By consequence, (P) and (C) are equivalent: the optimal values of both programs coincide, and an optimal solution (\mathbf{x}^, X^*) to (C) encodes an optimal solution to (P) in the sense that \mathbf{x}^* is in the convex hull of optimal solutions to (P).*

Proof Obviously we only have to deal with the feasible case. Since $\text{feas}(C)$ is closed and convex, its linear image $\text{feas}^+(C)$ is closed. Hence, by (2) we have always $\text{feas}^+(P) \subseteq \text{feas}^+(C)$ and it suffices to show

$$\overline{\text{feas}^+(P)} \supseteq \text{feas}^+(C) \tag{9}$$

in order to establish (8). If (9) were wrong, then there were a point in $\text{feas}^+(C)$ that could be separated by a hyperplane from $\text{feas}^+(P)$. Now the infimum of any linear function over $\text{feas}^+(C)$ coincides with the infimum over $\text{feas}^+(P)$. Indeed, it is clear that the latter is never smaller than the former. On the other hand, if the latter is finite, then (7) applies (to the separation functional rather than to the given objective, of course), and the arguments in [2, Theorem 2.6]¹ apply: they show equality of the two infima, in contradiction to the properties of a separating hyperplane. \square

Note that boundedness of $\text{feas}^+(P)$ is not an issue here: of course, if (P) is compact, then $\text{feas}^+(P)$ is also compact, as the convex hull of the continuous image of a compact set [3]. However, despite of Proposition (2.2), the weak key condition (4) does still not suffice to prove (8) or (9). In fact, there are cases where strict inclusion $\overline{\text{feas}^+(P)} \subset \text{feas}^+(C)$ can hold even if $\text{feas}(P)$ is *finite*, so that $\text{feas}^+(P)$ is a polytope and therefore compact. We owe this observation to an anonymous referee of an earlier version pointing out an example that satisfies (4), but for which the optimal objective values of (P) and (C) differ.

Acknowledgments Thanks are due to the anonymous referees whose comments significantly improved our presentation.

References

1. Berman, A., Shaked-Monderer, N.: Completely Positive Matrices. World Scientific, Singapore (2003)
2. Burer, S.: On the copositive representation of binary and continuous nonconvex quadratic programs. Math. Program Ser. A **120**(2), 479–495 (2009)
3. Stoer, J., Witzgall, C.: Convexity and Optimization in Finite Dimensions, Part 1. Springer, Berlin (1970)

¹ The proof in [2] considers linear functions with a zero in the $(1,1)$ -position, but can trivially be extended to any linear function.